Metropolia Ammattikorkeakoulu
Tietotekniikan koulutusohjelma

Jaani-Markus Kamila

Paravirtualisointi Xen virtualisointialustalla

Insinöörityö 1.4.2011

Ohjaaja: System Manager Mikko Mallat
Ohjaava opettaja: yliopettaja Matti Puska

Metropolia Ammattikorkeakoulu INSINÖÖRITYÖN TIIVISTELMÄ

Tekijä Otsikko Sivumäärä Aika	Jaani-Markus Kamila Paravirtualisointi Xen hypervisorilla 41 sivua 1.4.2011
Koulutusohjelma	tietotekniikka
Tutkinto	insinööri (AMK)
Ohjaaja Ohjaava opettaja	system manager Mikko Mallat yliopettaja Matti Puska

Tässä työssä tutkittiin yleisesti Linux-jakeluiden paravirtualisointia Xen virtualisointialustalla. Työn tavoitteena oli hyödyntää yrityksen palvelininfrastruktuuria siten, että jo olemassa olevia palvelinlaitteita hyödynnetään mahdollisimman tehokkaasti. Tehokkaan laitteistojen hyödyntämisen lisäksi tavoitteena oli myös laskea fyysisistä palvelimista aiheutuvia kuluja laskemalla niiden määrää. Olennaisena osana luodulle ympäristölle oli vikasietoisuus ja virhetilanteista toipuminen mahdollisimman nopeasti.

Työn teoriaosuudessa tutkittiin paravirtualisointia Xen virtualisointialustalla ja miten paravirtualisointi teoriassa toimii. Lisäksi esitellään muita virtualisointitekniikoita ja miksi valitsimme juuri Xen hypervisorin virtualisointituotteeksemme

Käytännön osuudessa rakennettiin vikasietoinen ympäristö, joka koostui virtuaalipalvelimia ajavista isäntäkoneista, verkkoyhteyksistä ja jaetuista levyresursseista. Kaikki ympäristön komponentit oli tarkoitus kahdentaa ja näin yhden komponentin, kytkimen tai fyysisen palvelimen vikatilanne ei näy asiakkaan suuntaan kuin pienenä katkoksena.

Hakusanat	Xen, hypervisor, virtualisointi, linux, paravirtualisointi, virtualisointialusta

Metropolia Polytechnic ABSTRACT

Author Name of Thesis Pages Date	Jaani-Markus Kamila Paravirtualization with Xen hypervisor 41 pages 1.4.2011
Degree Programme	Information Technology
Degree	Bachelor of Engineering
Instructor Supervisor	Mikko Mallat, System Manager Matti Puska, Principal Lecturer

The purpose of this Bachelor' thesis was to study paravirtualization with Xen hypervisor. The aim was to re-use servers we already had in production with maximum efficiency. Besides utilizing servers to the maximum level the goal was also to reduce physical server numbers and also cut running costs. It was also essential to create system which can recover from failures with almost no downtime.

The theoretical part of this thesis studied the concept of Xen paravirtualization and how it works in theory. In addition to Xen theory there is comparison to other virtualization techniques and why we choose Xen to be the product of choice.

The practical part of this thesis shows how we could create system which can quickly recover from failures. System consists of physical servers running virtual servers, network connections and shared disk storage. All components within the system was to be duplicated so no one component, switch of physical server failure could cause nothing but brief interruption of service.

Keywords	Xen, hypervisor, virtualization, linux, paravirtualization

Sisällys

Lyhenteet

chroot	Komento, jolla voidaan vaihtaa Linuxin käytössä olevaa root (”/”) -polkua.
DRBD	Distributed Replicated Block Device, mahdollistaa tiedostojärjestelmän replikoimisen useamman palvelimen kesken saavuttaen samalla korkean käytettävyyden.
EXT3	Journaloitu Linuxin yleisesti käyttämä tiedostojärjestelmä.
HVM	Hardware Virtual Machine tarjoaa uusia suoritustiloja, jossa virtuaalipalvelimia voidaan ajaa.
iSCSI	Internet Small Computer System Interface, IP-pohjainen standardi, jolla levypintaa voidaan jakaa lähettämällä SCSI-komentoja IP-verkossa.
LVM	Logical Volume Manager, jonka avulla tallennuslaitteet näkyvät abstraktimpana kokonaisuutena. LVM tarjoaa joustavamman tavan käsitellä tallennuskapasiteettia.
Nagios	Valvontaohjelmisto, joka valvoo palvelimia ja palveluita.
NFS	Network File system. Tapa jakaa levyosioita verkossa.
Rypäs	Useampi kuin yksi tietokone pyrkii luomaan yhteisistä resursseistaan yhden yhteisen resurssin. Tarkoituksena parantaa esimerkiksi saatavuutta tai suorituskykyä.
SCSI	Small Computer System Interface, standardi tiedon välittämiseksi tietokoneen ja sen oheislaitteiden välillä.
SMS	Short Message Service, tekstiviesti.
Stacking	Kytkimet, jotka on liitetty toisiinsa stacking-tekniikalla, saavat yhden kytkimen piirteet, mutta kaikkien kytkinten porttikapasiteetin.
Thin-client	Yksinkertainen tietokone, jolla on tarkoitus ottaa yhteys huomattavasti tehokkaampaan keskustietokoneeseen.
VNC	Virtual Network Computing on protokolla tietokoneen graafisen käyttöliittymän etäkäyttöön.

1 Johdanto

Työssäni eräällä pohjoismaisella palveluntarjoajalla on vuosien kuluessa laitekanta kasvanut lähes hallitsemattoman suureksi. Oli alettava etsiä ratkaisua, jolla ongelma tilojen, sähkönkäytön ja ilmastoinnin kanssa saataisiin hallintaan.

Palvelinten virtualisointi tuli puheeksi keväällä 2007. Siitä lähtien se on ollut vain ajatus paperilla, joten päätin ottaa projektin hoitaakseni. Lähes olemattomalla budjetilla oli hyödynnettävä jo olemassa olevia laitteita tavalla tai toisella. Myös ohjelmistoissa oli säästettävä. Päätinkin ryhtyä hakemaan ratkaisua avoimen lähdekoodin ohjelmistoista ja uhrata rahan sijaan enemmän aikaa järjestelmän toteuttamiseen.

Tavoitteenani oli rakentaa palvelinjärjestelmä, joka toteutettaisiin olemassa olevilla laitteilla hyödyntäen niitä aikaisempaa tehokkaammin. Saatavuustasoa pyrittiin parantamaan virtualisoinnin kautta niin, että palvelimet eivät ole riippuvaisia alla olevasta laitteistosta vaan ne sijaitsevat verkossa jaetuilla resursseilla. Yksittäisen komponentin tai fyysisen palvelimen virhetilanne ei siis aiheuta virtuaalipalvelimilla ajettaviin palveluihin virhetilaa.

Muutamasta vaihtoehdosta valitsin Xen hypervisorin (myöh. Xen) avoimen lähdekoodin version, jolla lupaillaan virtuaalipalvelinten saavuttavan lähes laitetason nopeuksia. Isäntäkoneiden käyttöjärjestelmänä käytän Gentoo Linuxia (versio 2007.0), mutta pyrin minimoimaan jakeluiden tuomat erot työssäni.

Työssäni pyrin luomaan palvelinjärjestelmän, joka pystyy hyödyntämään olemassa olevaa laitteistoa huomattavasti nykyistä tehokkaammin. Järjestelmän luotettavuus on myös keskeinen tavoite eikä järjestelmän minkään yksittäisen komponentin tai palvelimen virhetila saa aiheuttaa näkyvää haittaa ja tilanteesta on pystyttävä toipumaan nopeasti.

Virtualisointi on useimmille yrityksessäni työskentelevistä vielä melko vieras käsite. Työn tavoitteena onkin myös perehdyttää yrityksen teknistä henkilökuntaa virtualisointiin ja sen tuomiin etuihin sekä opastaa heidät tarvittavien apuvälineiden ja tulevan infrastruktuurin käyttöön ja hallintaan.

Tässä työssä tutkitaan paravirtualisointia, suorituskykyä ja saatavuutta. Työn lopussa kuvataan, miten yksinkertainen palvelin paravirtualisoidaan. Työ ei käsittele verkon topologiaa tai iSCSI:n (Internet Small Computer System Interface) käyttöön tarvittavien asennusten tekemistä. Myöskään HVM-virtualisointia (Harware Virtual Machine) työ ei kata, koska käytössäni ei ollut siihen kykeneviä laitteita, eikä sillä olisi saavutettu Linuxin virtualisoinnin kohdalla mainittavia etuja.

2 Virtualisointi

Viime vuosina palvelinten suorituskyky on kasvanut niin nopeasti, että useat tehopalvelimet eivät käytä resursseistaan kuin pienen murto-osan. Samaan aikaan kuitenkin palvelinten lukumäärä on kasvanut jatkuvasti kiihtyvällä vauhdilla. Lukumäärän kasvu on ollut selitettävissä sillä, että erillisillä palvelimilla saavutetaan hyötyä niin tietoturvassa, virhetilanteidenkin eristyksessä kuin komponenttien mitoituksessakin. Myöskään asiakas ei useinkaan halua jakaa ympäristöään muiden kesken vaan oma palvelin on ainoa vaihtoehto. (3.)

Virtualisointi on ratkaisu kahden suuntauksen – keskitettyjen ja hajautettujen järjestelmien välillä. Sen sijaan, että hankitaan ja ylläpidetään kokonainen uusi fyysinen palvelin komponentteineen, voidaan sovelluksia ajaa virtuaalisessa ympäristössä. Virtuaaliselle ympäristölle annetaan käytettävissä olevista resursseista tietty prosessoriteho, muisti ja levypinta. Näin vapaista resursseista luotu ympäristö saavuttaa samat ominaisuudet kuin fyysinenkin palvelin, kuitenkin jakaen fyysisen palvelimen resursseja tehokkaasti. (3.)

Virtualisointi murtaa yleisen käsityksen tietokoneen arkkitehtuurista, erottaen käyttöjärjestelmän fyysisestä alustasta, jolla tätä ajetaan. Näin mahdollistetaan useiden käyttöjärjestelmien ajo samalla fyysisellä alustalla niin, että ne eivät ’tiedä’ toistensa olemassaolosta tai pysty vaikuttamaan toisiinsa. Tällainen toteutus myös estää virhetilanteiden leviämisen virtuaalipalvelimelta toiselle. Jokaisella luodulla virtuaalipalvelimella on käytössään sille allokoidut resurssit, joilla ohjelmia suoritetaan aivan kuten fyysiselläkin palvelimella. Näitä resursseja virtualisointialusta jakaa virtuaalipalvelinten kesken. (3.)

2.1 Virtualisoinnin hyödyt

Yrityksiä virtualisointiin on johtanut ajan myötä kertyneet palvelinlaitteet, jotka suorittavat jotakin tiettyä tehtävää yrityksen toiminnassa. Useassa erillisessä palvelimessa ajetut palvelut voidaan kuitenkin helposti siirtää samalle isäntäkoneelle virtuaalipalvelimiksi. Esimerkiksi on mahdollista tarjota Sendmailia FreeBSD:llä ja samalla Apachen web-palvelinta Red Hat Enterprisellä, molemmat samalla fyysisellä palvelimella ajettuina. Tämä puolestaan laskee kuluja niin energian kuin tilojen ja ilmanvaihdonkin puolesta. Myös palvelinten hallinta helpottuu huomattavasti, kun kaikki ajossa olevat virtuaalipalvelimet ovat hallittavissa yhdestä näkymästä. Virtualisointi tarjoaa myös paremmat mahdollisuudet pitää palvelimia tavoitettavissa vuorokauden ympäri. Esimerkiksi häiriötilanteessa tai muuten alas ajettu palvelin voidaan korvata sitä vastaavalla virtuaalipalvelimella. Myös virtuaaliset ryppäät ovat laajalti nykypäivänä käytössä olevia ratkaisuja. (2.)

Merkittävä etu virtualisoinnissa on myös vanhempien fyysisten palvelinten yhdistäminen tai siirto uudelle alustalle. Palvelinten tekniikka on vuosien saatossa kehittynyt niin paljon, että vanhemmat järjestelmät eivät välttämättä ole yhteensopivia uuden tekniikan kanssa. Tämä johtaa väistämättä ongelmiin laitteiston uusimisen yhteydessä ja esimerkiksi virhetilanteessa, jolloin vanhempia varaosia ei välttämättä ole kovinkaan helposti saatavilla tai vanhempaa ohjelmistoa ei ole mahdollista ajaa uudemmalla alustalla. Virtualisointi tarjoaa laajasti yhteensopivan alustan, jolla vanhempiakin järjestelmiä pysytään tarjoamaan varsin kitkattomasti.

Menneinä vuosina käyttöjärjestelmät olivat sidottuja tiettyyn arkkitehtuuriin. Nykypäivänä kuitenkin useimpia käyttöjärjestelmiä löytyy laajalle arkkitehtuurien skaalalle, joista suosituimpana x86, jolla voidaan ajaa muun muassa Windowsia, UNIXia tai vaikkapa joitakin Linux-jakeluita, muutamia mainitakseni. x86-virtualisointitekniikat voivat vuorostaan taas toimia isäntinä näille ympäristöille, jolloin käyttöjärjestelmiä voidaan ajaa virtualisoituina saman arkkitehtuurin tarjoavalla isäntäkoneella. (3.)

Palvelinten virtualisoinnilla voidaan tehokkaan resurssien hyödyntämisen lisäksi säädellä helposti, millä isäntäkoneella tiettyjä virtuaalikoneita ajetaan. Näin resurssien jakamista on vielä helpompaa entisestään tehostaa. Esimerkkinä voisi olla vaikkapa automatisoitu järjestelmä, joka kuormitustilanteessa siirtää osan ajettavanaan olevista virtuaalipalvelimista toiselle isäntäjärjestelmälle.

Kuva 1. Thin-client ja pöytätietokone

Tänä päivänä yksittäiset työpöytäkoneet ovat usein yhdistelmä käyttöjärjestelmää, ohjelmistoja ja käyttäjän asetuksia. Näitä yksittäisiä työpisteitä hallitaan yleisesti kankeasti ja hitaasti yksi kerrallaan. Monet organisaatiot ovatkin kääntäneet katseensa työpöytien virtualisointiin. Virtuaaliset työpöydät korvaisivat vanhan ja kankean lähestymistavan joustavammalla ratkaisulla. Ylläpitäjät voisivat rakentaa erilaisiin tarkoituksiin valmiita käyttöjärjestelmäpaketteja, joita sitten hallittaisiin keskitetysti. Käyttäjälle annettaisiin thin-client (kuva 1), jolla tehopalvelimella ajettavaa virtuaalityöpöytää käytettäisiin samaan tapaan kuin lokaalia käyttöjärjestelmää. Uusin teknologia antaa mahdollisuudet jopa virtuaalisen työpöydän ottamisen mukaan esimerkiksi kannettavalla tietokoneella, kunhan levynkuva vain kopioidaan palvelimelta tietokoneelle, jolla sitä halutaan ajaa. Myöhemmin levynkuva voidaan synkronoida palvelimen kanssa. (9.)

2.2 Virtualisointitekniikoita

Virtualisointiohjelmistoista suurin osa esittelee virtuaalipalvelimille täysin tai osaksi virtualisoidun palvelinlaitteiston, mutta osa tuotteista pystyy osoittamaan virtuaalipalvelimille myös fyysisiä palvelimen osia kuten PCI-portin (Peripheral Component Interchange). Kaikkien virtualisointiohjelmistojen tavoite on ajaa virtualisoituja ympäristöjä luotettavasti ja toisistaan eristetysti hyödyntäen mahdollisimman tehokkaasti fyysistä palvelinlaitetta. Enemmistö näistä ohjelmistoista on suunnattu yrityskäyttöön. Muutamia tunnetuimpia ovat VMware ESX, Microsoft Virtual Server ja XenSourcen XenServer (kaupallinen Xenin versio). (3.)

Taulukko 1. Virtualisointituotteiden vertailu

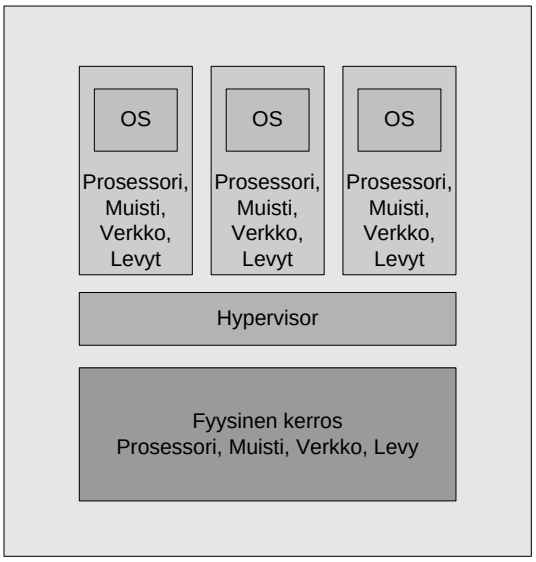
Nimi	Kehittäjä	Arkkitehtuuri	Virtualisoidut käyttöjärjestelmät	SMP	Kaupallinen tuki
KVM	KVM Project	x86	Linux, Windows	Ei	Ei
Virtual Iron	Virtual Iron Software Inc.	x86	BSD, Linux, OS/2, Windows	Kyllä	Kyllä
VMware ESX	VMware	x86	FreeBSD, Netware, Solaris, Windows	Kyllä	Kyllä
Virtual Server	Microsoft	x86	Linux, Windows	Kyllä	Kyllä
Xen	Cambridgen yliopisto	x86	xBSD, Linux, Solaris, Windows	Kyllä	Ei
XenServer	XenSource	x86	Linux,Windows	Kyllä	Kyllä

Valitsin Xenin virtualisointituotteeksi, koska se oli työni aloituksen aikaan yksi tunnetuimmista ja käytetyimmistä virtualisointiohjelmistoista, joka perustuu avoimeen lähdekoodiin. Xen oli myös avoimen lähdekoodinsa ansiosta ilmainen, ja se linjasi hyvin erittäin pienen budjettini kanssa. Yllätykseksenі Xen oli kuitenkin saavuttanut laajan käyttäjäkunnan, joten viitseliäs ihminen löysi tukea laajalti vain hakemalla Internetistä. Kaupallista tukea ei projektini alkuvaiheessa ollut saatavilla, mutta nykyään XenSource tarjoaa kaupallista versiota Xenistä tuotenimellä XenServer. Taulukossa 1 voidaan nähdä erilaisten virtualisointituotteiden vertailua. (3.)

2.3 Xen virtualisointialusta

Vaikka on useita tapoja virtualisoida resurssit käyttäen ns. virtualisointialustaa, jota Xenin tapauksessa kutsutaan nimellä hypervisoriksi, on näillä kaikilla sama toimintaperiaate. Käyttöjärjestelmää ajetaan itsenäisesti ja eristetysti samaan tapaan, miten se toimisi myös fyysisellä palvelimella. (3.)

Virtualisointialusta tai hypervisor tarkoittaa kerrosta käyttöjärjestelmän ja fyysisen rajapinnan välissä. Tämän kerroksen päällä voidaan ajaa yhtä tai useampaa virtuaalista käyttöjärjestelmää, joita hypervisor valvoo.



Kuva 2. Hypervisorin sijoittuminen

Hypervisor on Xen-virtualisoinnin ydin, joka hallitsee fyysisiä resursseja ja pyyntöjä virtuaalisilta käyttöjärjestelmiltä resursseille. Se osoittaa virtuaalisen prosessorin (tai useampia), muistin, I/O- ja levyresursseja alaisuudessaan oleville virtuaalipalvelimille (kuva 2).

Tämän työn pääasiallista lähestymistapaa virtualisointiin kutsutaan paravirtualisoinniksi.

Paravirtualisoinnilla tarkoitetaan sitä, että virtualisoitavaa käyttöjärjestelmää muunnellaan niin, että käyttöjärjestelmä ohjaa laitteistolle tarkoitetut operaatiot virtualisointialustalle. Virtualisointialusta on puolestaan vastuussa suorasta kosketuksesta isäntäkoneen laitteistoon ja näin säätelee virtualisoitujen käyttöjärjestelmien resursseja. (5.)

Xenillä paravirtualisoidun virtuaalipalvelimen vaatimukset:

- Linux-ydin, johon on käännetty tuki virtualisointialustalle.
- Tiedostojärjestelmä, joka sisältää virtuaalipalvelimen tarvitsemat tiedostot.
- Ytimen moduulit, joita ajettava virtuaalipalvelin tarvitsee. Näitä ei välttämättä tarvita jos tarpeelliset moduulit esimerkiksi käännetään staattisesti käytettävään ytimeen.
- Swap-osio tai -tiedosto.
- Xen konfiguraatiotiedosto.

Tyypillisesti Linux-ydin sijaitsee tiedostojärjestelmässä /boot -hakemistossa ja moduulit /lib/modules -hakemistossa. Tämä ei kuitenkaan pidä paikkaansa paravirtualisoidun Linuxin kohdalla. Virtuaalipalvelimen tarvitsema ydin sijaitsee isäntäkoneella valinnaisessa paikassa, josta se käynnistyksen yhteydessä tarjotaan virtuaalipalvelimelle. Tämä esimerkiksi estää virtuaalipalvelimella ytimen uudelleen kääntämisen, ja se onkin aina tehtävä isäntäkoneella. Moduulit sen sijaan on kopioitava virtuaalipalvelimelle kuten fyysisellekin palvelimelle, eli /lib/modules -hakemistoon. Tämä on yksi syy kääntää moduulit suoraan staattisesti ytimeen, jotta säästytään moduulien kopioimiselta päivitysten yhteydessä.

Täysin virtualisoidun (HVM) palvelimen vaatimukset ovat

- levytila root- ja swap-osioille
- Xen-konfiguraatiotiedosto
- asennusmedia, jolta täysin virtualisoitu palvelin asennetaan.

Jos vapaana ei ole fyysistä levypintaa, kuten osioita tai logista asemaa (LVM), niin täysin virtualisoidulle palvelimelle voidaan osoittaa myös tiedostopohjainen tiedostojärjestelmä. Tiedostopohjaisella tiedostojärjestelmällä tarkoitetaan levynkuvan luomista isäntäpalvelimen tiedostojärjestelmään, jonka sisään luodaan uusi tiedostojärjestelmä.

Levynkuva tiedostojärjestelmää varten voidaan luoda esimerkiksi dd-komennolla, joka luo halutun kokoisen tiedoston. Luodun tiedoston sisään voidaan alustaa tiedostojärjestelmä samaan tapaan kuten fyysiselle kovalevylle. Tähän tiedostoon on tarkoitus asentaa valittu Linux-jakelu.

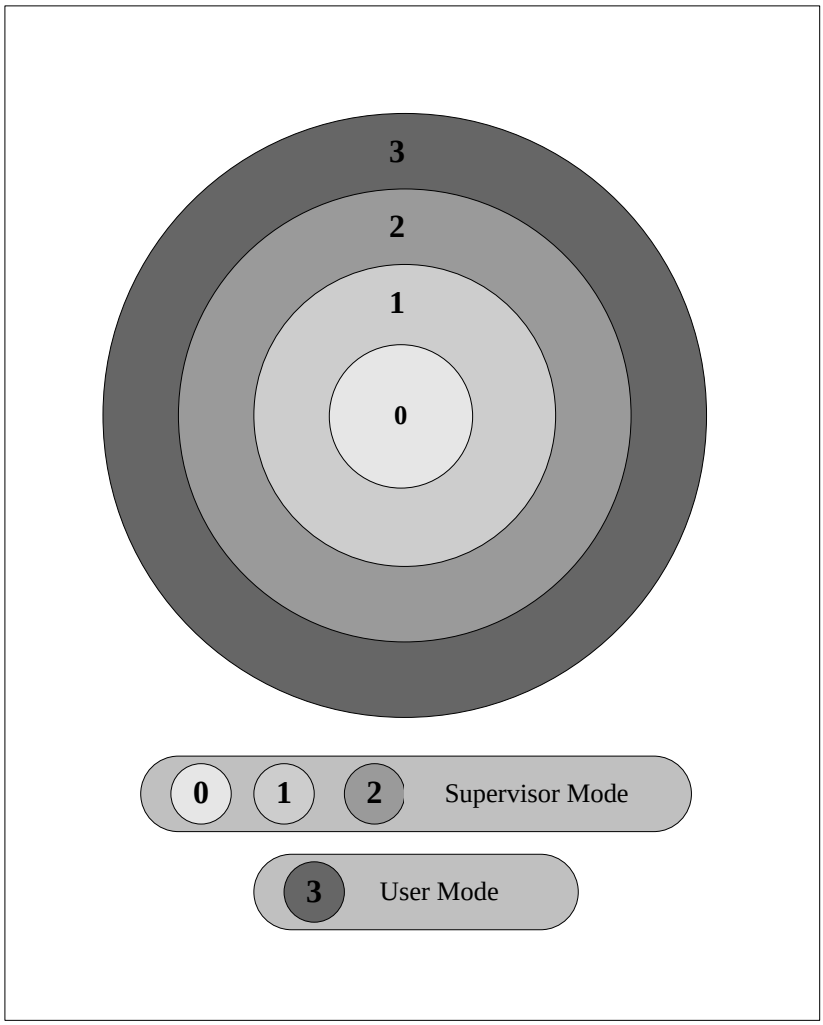
Täysin virtualisoidun palvelimen ajoon tarvitaan lisäksi joko Intelin VT- tai AMD:n AMD-v -prosessori. Nämä prosessorit sisältävät tuen HVM-virtualisoinnille. Prosessorin HVM-tuen voi tarkastaa prosessorin ominaisuusluettelosta. Nykyaikaisissa palvelimissa tuki on jo lähes poikkeuksetta.

Hypervisoreita on kahta tyyppiä:

- *Tyypin 1 hypervisor:* Virtualisointiohjelmistoa (hypervisor) ajetaan suoraan fyysisellä alustalla, jota kutsutaan nimellä Ring-0. Näin ollen sen hallinnassa olevia virtuaalipalvelimia ajetaan virtualisointialustasta katsottuna seuraavalla tasolla (Ring-1) kuten kuvasta 3 nähdään. Xen kuuluu tähän ryhmään.
- *Tyypin 2 hypervisor:* Virtualisointiohjelmistoa ajetaan käyttöjärjestelmän päältä, useimmiten Ring-3:ssa, joka on korkein suoritustaso x86-arkkitehtuurissa. Tämä tarkoittaa sitä, että kun virtuaalipalvelimet ovat näinkin kaukana fyysisistä resursseista, vähenee niiden suorituskyky huomattavasti. Suorituskyvyn väheneminen johtuu siitä, että ohjelmiston kutsut fyysisille resursseille joutuvat kulkemaan useamman kerroksen läpi päästäkseen perille. Tähän ryhmään kuuluvat esimerkiksi VMware Workstation- ja VirtualBox-ohjelmistot.

Käytetyin prosessoriarkkitehtuuri nykyään on x86-yhteensopiva. Alkaen 80286-piirisarjasta se tarjosi kaksi tapaa osoittaa muistia: avotila ja suojatila. Myöhemmin 80386:sta eteenpäin mukaan tuli kolmas tapa nimeltään virtuaalitila. Tämä mahdollisti avotilalle kirjoitettujen ohjelmien ajon kuitenkin kiertäen avotilan rajoitteet tarkempiin yksityiskohtiin menemättä. Näin ohjelmia ei välttämättä tarvinut viedä suojatilaan. Avotila, joka on rajoitettu vain yhteen megatavuun muistia, jäi nopeasti tarpeettomaksi ja virtuaalitila, joka toimi vain 16-bittisissä järjestelmässä putosi 32-bittisten käyttöjärjestelmien yleistyttyä x86-arkkitehtuurille. Sittemmin suojatila tarjosi useita uusia ominaisuuksia, kuten tuen moniajolle ja prosessien segmentoinnin. Segmentoinnin myötä prosessit eivät voineet enää kirjoittaa oman muistialueensa ulkopuolelle. Suojatila tarjosi myös laitteistotason tuen virtuaalimuistille. x86-perheen suojatila käyttää neljää oikeustasoa, tai kehää (ring), numeroituna nollasta kolmeen. Käytännössä siis muisti ositetaan, jaetaan segmenttien kesken ja nämä segmentit kiinnitetään tiettyyn kehään. Termi "Ring" juontaa juurensa GE:n ja MIT:n rakentamaan järjestelmään, jolla eri oikeustasoja visualisoitiin sisäkkäisillä renkailla. Ring-0 tarkoittaa sisäisintä rengasta, jolla on täydellinen kontrolli prosessoriin. Ring-3 puolestaan on ulommaisin kehä ja oikeuksiltaan rajoittunein. (3.)

Kuvassa 3 nähdään x86-arkkitehtuurin oikeustasot



Kuva 3 x86-arkkitehtuurin oikeustasot

Esimerkiksi VMware tarjoaa kaupallisia ratkaisuja, kuten ESX Server ja VMware Workstation. Ohjelmisto, kuten VMware Workstation, käyttää ns. täyttä virtualisointia. Se virtualisoi tietokoneen kaikki piirteet. Tämänkaltainen virtualisointi johtaa suurempiin viiveisiin virtualikoneiden ja isäntäkoneen tarjoamien resurssien välillä, jolloin samanaikaisesti ajetut käyttöjärjestelmät toimivat hitaammin kuin tavallisesti. Komponenttien hinnat kuitenkin laskevat, ja ohjelmistojen koodia optimoidaan kokoajan. Tästäkin ongelmasta saatetaan mahdollisesti tulevaisuudessa päästä eroon. Tällä hetkellä se kuitenkin on huomattava ongelma. Kannattaa kuitenkin huomioida, että esimerkiksi ESX Server tarjoaa parempaa suorituskykyä kuin Workstation versio,

ja se onkin tarkoitettu yrityskäyttöön. ESX on laajalti käytetty virtualisointiratkaisu Xenin ohella. (3.)

VMwaren tuotteiden etu on ollut sen tarjoama mahdollisuus ajaa täysin muuntelemattomia käyttöjärjestelmiä, mikä on ollut Xenillä mahdollista vasta versiosta 3.0.0 lähtien. Tämä tarkoittaa sitä, että kun järjestelmä paravirtualisoidaan Xenillä, niin ajettavaa käyttöjärjestelmää on myös muokattava. Linuxin tapauksessa vaatimuksena on kuitenkin vain Xen-tuki ytimeen ja nykyiset Linux-jakelut tarjoavat tämän useimmiten valmiina. Windowsin kohdalla paravirtualisointi on vielä tällä hetkellä mahdotonta. Täysin virtualisoidut (HVM) käyttöjärjestelmät puolestaan toimivat ilman muuntelua. (11.)

Tämä ei kuitenkaan tarkoita sitä, että täysin virtualisoidut käyttöjärjestelmät toimisivat aivan ongelmitta Xen-virtualisointialustalla. Esimerkiksi Microsoft Windowsin suorituskyky on huomattavan heikko ja isäntäkoneessa vaaditaan Intelin tai AMD:n virtualisointia tukeva prosessori Intel-VT- tai AMD-V-sarjasta. Microsoft Windowsin suorituskykyyn löytyy parannuksia XenSource Inc. -nimisen yhtiön kehittämistä ajureista, joilla päästään melko hyviin tuloksiin. XenSource Inc. tarjoaa myös kaupallista versiota Xenistä. Täysi virtualisointi on kuitenkin Xenin tapauksessa ainoa keino ajaa Microsoft Windowsia. (11.)

3 Ympäristön kuvaus

Projektin tavoitteena oli rakentaa vikasietoinen ja olemassa olevia resursseja tehokkaasti hyödyntävä ympäristö mahdollisimman pienillä kustannuksilla. Lisätavoitteena voidaan pitää myös hintojen laskua tulevaisuudessa, koska esimerkiksi sähköä kuluu vähemmän eikä ilmanvaihtoon tarvita enää niin suuria investointeja kuin ennen.

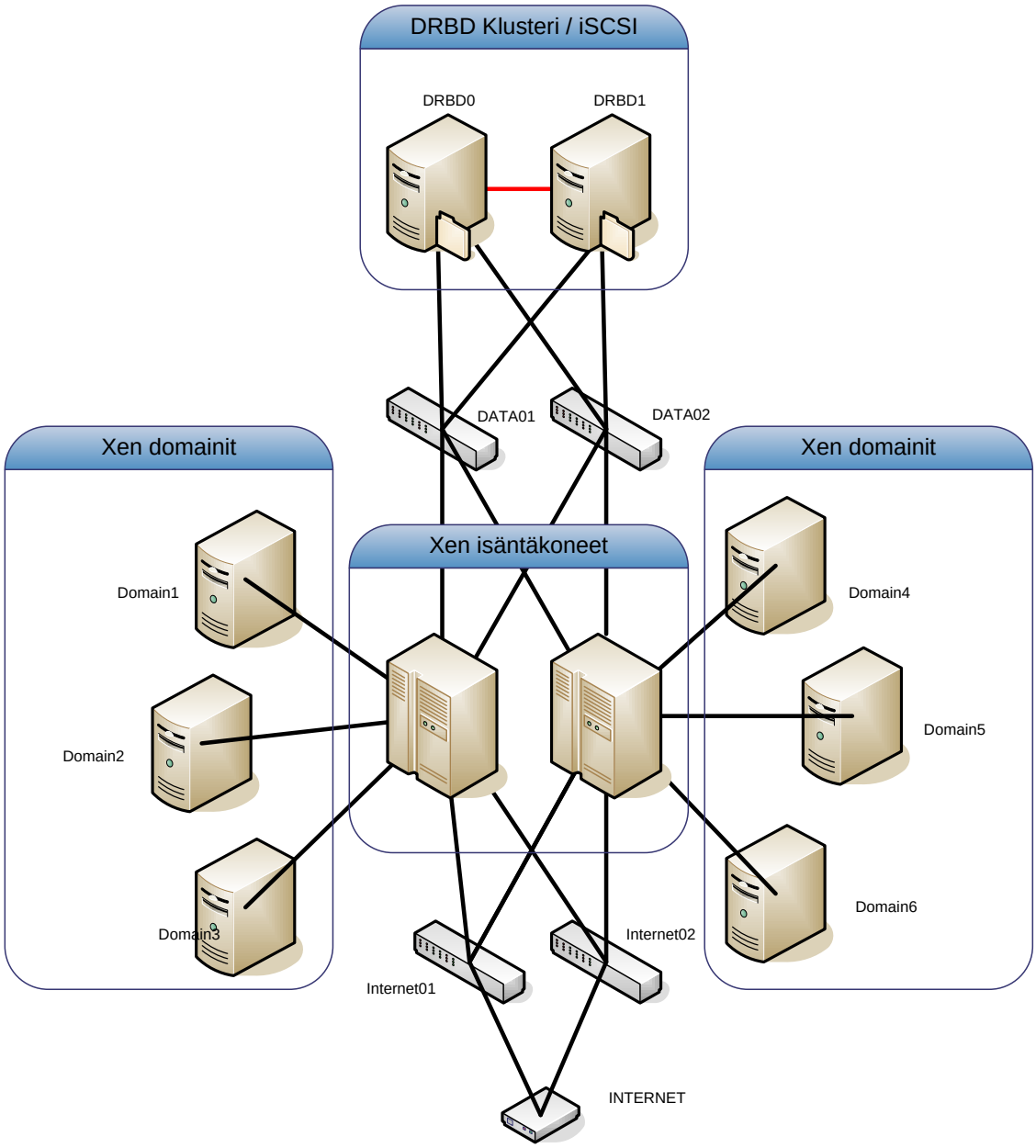
Vikasietoisuus on verkkokerroksessa toteutettu kahdentamalla kaikki verkkolaitteet niin, että yhden laitteen vikatilanne ei aiheuta näkyvää virhettä verkkoon. Levypinta on myös kahdennettu rypäsratkaisulla, josta virtuaalipalvelimet saavat niille allokoidun levypinnan käyttöönsä verkon avulla. Rypäs koostuu kahdesta isosta levypalvelimesta, joista toinen voi olla milloin tahansa tavoittamattomissa ilman, että se näkyisi mitenkään sitä hyödyntäville virtuaalipalvelimille.

Isäntäpalvelimiin on myös laskettu sen verran muistikapasiteettia, että yksi isäntäpalvelin voi vikaantua ja sen ajamat virtuaalipalvelimet saadaan ajoon jollakin toisella isäntäpalvelimella hyvinkin nopeasti.

Luodussa ympäristössä oli tarkoitus kahdentaa kaikki kriittiset komponentit ja näin taata luotettava käytettävyys ja palveluiden toiminta.

4 Palvelutaso

Tunnetusti palveluiden on oltava tarvittaessa aina käytettävissä. Tästä onkin tullut useille yrityksille suuri huolen aihe, koska sillä on melko suora suhde myös voiton tuottamiseen. Yritykset ovatkin usein investoineet suuria summia palvelin-infrastruktuuriinsa ja siten pyrkineet varmistamaan, että kriittiset palvelut ovat aina saatavilla ja toiminta voi jatkua vaikeissakin olosuhteissa katkeamatta. Investoinnit varajärjestelmiin ovat osoittautuneet kalliiksi ratkaisuksi. (3.)

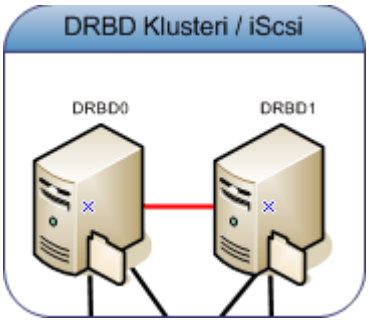


Kuva 4 Verkko

4.1 Saatavuus

Mahdollisimman luotettavan saatavuuden takaamiseksi olen työssäni pyrkinyt kahdentamaan kaiken mahdollisen, kuten verkkoyhteydet, levypinnan ja Xen-isäntäkoneet, joilla varsinaisia virtuaalipalvelimia ja siten palveluja ajetaan. Kuvassa 4 voidaan nähdä järjestelmän kahdennus ja yksittäiset komponentit.

Levyt



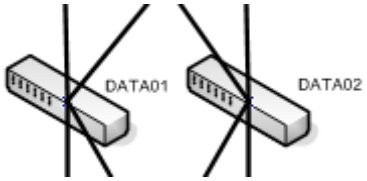
Kuva 5 Levypalvelin

Levypalvelimen (kuvassa 5):

- HP Proliant DL380 G3
- 2 x 2.4 GHz P4 Prosessori
- 2 GB RAM
- 6 x 72 GB 15krpm SCSI-levyt
- RAID-5.

Levypintaa virtuaalipalvelimille tarjotaan iSCSI- ja LVM-tekniikoilla (Logical Volume Manager), joita tässä työssä käsitellään vain hyvin pintapuolisesti. iSCSI on tekniikka, jolla SCSI-komentoja voidaan lähettää verkon avulla ja kyseiseen verkkoon liitetyt isäntäpalvelimet voivat ottaa nämä levyalueet käyttöönsä tarvittaessa. Tämä antaa virtuaalipalvelimille mahdollisuuden ns. live migraatioon, jota käsitellään luvussa 5.2.3 yksityiskohtaisemmin. Samalla virtuaalipalvelinta voidaan ajaa millä tahansa isäntäkoneella kunhan sillä vain on verkkoyhteys ja tarvittava levyalue saatavillaan. iSCSI:n toiminta on tämän työn virtuaalipalvelimille tärkein osa infrastruktuuria, joten se on pyritty kahdentamaan kaikin mahdollisin keinoin. Kyseistä palvelua tarjotaankin verkkoon DRBD-ryppäältä (Distributed Replicated Block Device), joka käsittää kaksi identtistä tehopalvelinta ja hyvin suuren kapasiteetin ja suorituskyvyn fyysistä levypintaa. Kahdennettu DRBD-rypäs tarkoittaa sitä, että toinen ryppään palvelimista voi olla milloin tahansa tavoittamattomissa ja iSCSI toimii normaalisti. Käytössä on aktiivi-passiivi mallin rypäs, jolloin toinen ryppään koneista käsittelee pyyntöjä ja toinen on varalla, aktivoituen virhetilanteen sattuessa. (12.)

Verkko



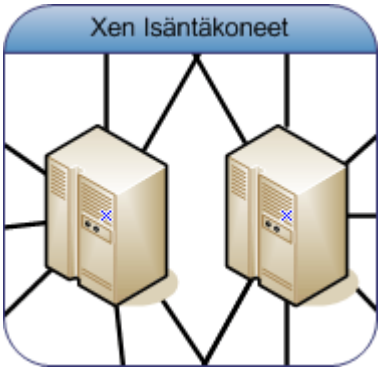
Kuva 6 Data-kytkimet

Kytkinten (kuvassa 6) tiedot ovat seuraavat:

- HP2424 24-porttinen kytkin
- 10 Mbit/s / 100 Mbit/s

Koska verkkoliitännät näyttelevät myös erittäin tärkeää roolia verkossa jaettavissa resursseissa, on näistä jokainen kahdennettu käyttäen aktiivi-passiivi -sidontatekniikkaa. Aktiivi-passiivi sidonnalla tarkoitetaan sitä, että kaksi verkkoliitäntää toimii koko ajan toisesta siihen liitetystä verkkoliitännästä, jolloin näistäkin toinen voi aina olla tavoittamattomissa häiritsemättä verkon toimintaa. Toimiakseen luotettavasti tämä tarvitsee rinnalleen myös kahdennetut kytkimet, jotka on kytketty toisiinsa pino-tekniikalla. Näistäkin toinen voi olla milloin tahansa esimerkiksi huollossa häiritsemättä verkon toimintaa. Teoriassa esimerkiksi verkkolaitteiden päivitykset ovat täysin mahdollisia ilman katkoja palveluissa.Verkon kuva on nähtävissä kuvassa 4.

Isäntäkoneet



Kuva 7 Isäntäkoneet

Isäntäkoneiden (kuvassa 7) tiedot ovat seuraavat:

- HP Proliant DL380 G3
- 2 x 2.4 GHz P4 Prosessori
- 8 GB RAM
- 2 x 18 GB 10 krpm SCSI-levyt
- RAID-1.

Isäntäkoneiden kapasiteetti on suunniteltu niin, että niistäkin vähintään yksi voi olla huollettavana ja virtuaalipalvelimia voidaan ajaa normaalisti jäljellä olevilla isäntäkoneilla.

Tässä kohtaa on kuitenkin mainittava yksi erittäin ikävä varjopuoli, joka koskettaa ainakin Xenin open source -versiota. Sisäänrakennettuna ei löydy mitään tapaa, jolla yhden isäntäkoneen vikaantuessa sen ajamat virtuaalipalvelimet siirrettäisiin automaattisesti muiden isäntäkoneiden ajoon. Tietysti tämä on mahdollista toteuttaa käyttäen omia tarkistuksia, joiden perusteella virtuaalipalvelin tarvittaessa käynnistetään. Tämä on kuitenkin hieman riskialtista, koska jos kaksi samaa virtuaalipalvelinta käynnistetään eri isäntäkoneilla samanaikaisesti, ne ottavat käyttöönsä saman levyalueen. Tämä aiheuttaa hyvin suurella todennäköisyydella tiedostojärjestelmän korruptoitumisen, koska käytössä olevaa tiedostojärjestelmää (EXT3) ei ole suunniteltu käytettäväksi samanaikaisesti useammalta palvelimelta luku-kirjoitustilassa. Virtuaalipalvelimet onkin siis käynnistettävä käsin isäntäkoneen vikaantuessa. Tästä ei kuitenkaan aiheudu suuria ongelmia, koska vikatilanteen sattuessa Nagios-valvontaohjelmisto lähettää viestin ylläpitäjille.

4.2 Suorituskyky

Virtuaalipalvelin on aina hieman fyysistä palvelinta hitaampi johtuen siitä, että virtualisointi edellyttää ylimääräisen kerroksen, virtualisointialustan, lisäämistä osalle suoritettavista operaatioista.

Tässä tapauksessa ei ollut järkevää kerätä testidataa ennen ja jälkeen virtualisointi-prosessin yksinkertaisesti siksi, että isäntäkoneet koostuivat hieman eri komponenteista kuin entiset fyysiset palvelimet. Myös fyysisen levypinnan allokointi tapahtuu tulevaisuutta silmällä pitäen LVM:n avulla, jolloin sekään ei ole vertailukelpoinen suoran fyysisen levyn allokointiin. LVM tarjoaa tulevaisuudessa mahdollisuuden laajentaa jo olemassa olevia osioita helposti ja minimaalisella käyttökatkolla.

Clarksonin yliopistossa on tehty suorituskykytestejä aivan Xenin alkuajoista lähtien. Näillä tutkimustuloksilla onkin ollut suuri merkitys siihen, miten Xen sai huomiota niin yritysten kuin akateemisten piirienkin keskuudessa. Suorituskyky on karkeasti noin 3,5% huonompi paravirtualisoidulla Linuxilla kuin natiivina ajettuna. (7.)

4.3 Tietoturva

Xen-asennuksissa on tärkeää suojata ensisijaisesti isäntäkone, koska murrettu isäntäkone tarjoaa pääsyn kaikkialle sen ajamille virtuaalipalvelimille. Muutamia hyviä käytäntöjä ovat seuraavat:

- Isäntäkoneella tulee ajaa mahdollisimman vähän palveluja. Mitä vähemmän palveluja tai prosesseja on käynnissä, sitä vähemmän mahdollisella hyökkääjällä on niin sanottua hyökkäyspinta-alaa
- Liikennettä isäntäkoneille tulee rajoittaa palomuurein
- Migraatioita tai muutakaan hallinnollista verkkoliikennettä ei tule sallia muualla kuin hallintaverkossa
- Käyttäjiä ei tule päästää isäntäkoneille, ainoastaan virtuaalipalvelimille. (1.)

Työn ympäristö on toteutettu niin, että jokaisessa virtuaalipalvelimessa on kaksi verkkokorttia. Toisella hoidetaan liikenne hallinta- ja dataverkossa, tosella taas yhteydet ulkomaailmaan, Internetiin.

Eriytettyä hallintaverkkoa pidetäänkin turvallisimpana verkkokonfiguraationa. Mielessä on kuitenkin pidettävä, että myös kaiken muun hallintaverkossa olevan on oltava yhtä lailla luotettua. (1.)

Erilaisiin tehtäviin kannattaa myös luoda erilaisia virtuaalipalvelimia, joilla on erilaiset vaatimukset tietoturvan suhteen. Esimerkkinä www-palvelin, jolla ajetaan vain tähän tarkoitukseen käytettyjä palveluja, ei mitään muuta. Palvelimen murron sattuessa hyökkääjä pääsee käsiksi mahdollisimman pieneen määrään tietoa, eikä esimerkiksi yrityksen sisäiseen tiedonjakoon, joka olisi myös hyvin voitu toteuttaa samalla palvelimella. Virtuaalipalvelimet tarjoavat näin eristetyt ympäristöt, ja uhkien hallinta helpottuu. (7.)

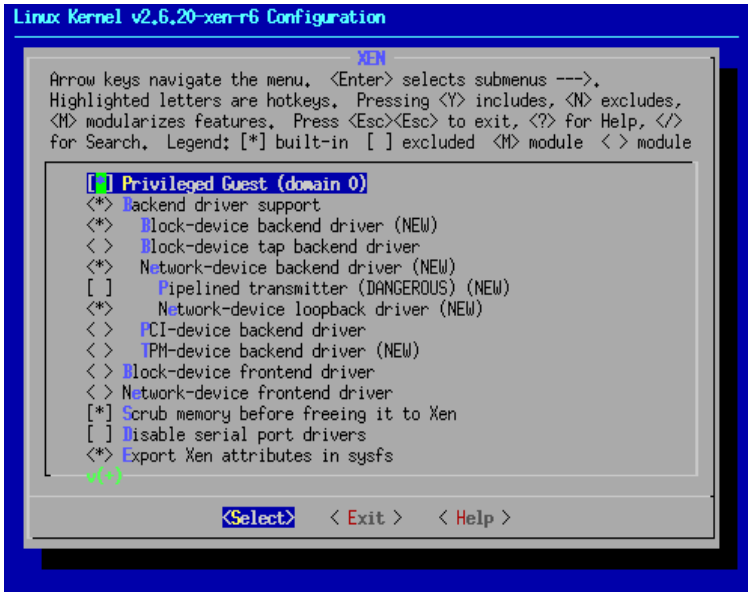
5 Xen virtualisointialusta

Kuten aikaisemmin todettiin, Linux tai muu käyttöjärjestelmä tarvitsee paravirtualisoituna toimiakseen muunnellun ytimen ja ajureita. Paravirtualisointiprosessia voidaan verrata siihen, että käyttöjärjestelmä siirretään fyysisestä koneesta toiseen. Alla oleva laitteisto muuttuu, joten myös ytimen ajurituki täytyy luoda uudelleen. Linux-järjestelmissä tämä käytännössä tarkoittaa sitä, että ydin tulee kääntää uudelleen tarvittavin uusin ajurein. Xenillä Linuxia virtualisoitaessa tarvitaan ytimeen Xen-ajurit. Kuitenkaan aikaisemmin asennettuihin ohjelmistoihin ei tarvitse tehdä mitään muutoksia, joten työmäärä on kuitenkin verrattain pieni. Useimmat nykyaikaiset Linux-jakelut sisältävät tarvittavan Xen-tuen oletusarvoisesti.

5.1 Käyttöönotto

Isäntäkone asennetaan kuten mikä tahansa Linux-jakelu. Ainoana erona on, että ydin vaihdetaan versioon, jossa on Xen-tuki. Tukea voidaan kutsua myös ns. tausta-ajurien tueksi, jotka puolestaan tarjoavat rajapinnan virtuaalipalvelinten käyttämille ns. etualan ajureille. Näiden ajureiden avulla virtuaalipalvelin ja isäntäkone kommunikoivat keskenään. Ero isäntäkoneen ja virtuaalipalvelimen ydinten konfiguraatioiden välillä ovat ainoastaan nämä ajurit. Suositus on myös poistaa SCSI-ajurit virtuaalipalvelimen ytimestä, koska ne saattavat joissain tapauksissa aiheuttaa virhetilanteita. Myöskään SCSI-ajureita ei tarvita mihinkään, koska fyysistä levyä ei virtuaalipalvelimille suoraan tarjota. Virtuaalipalvelimelle voidaan myös antaa suoraan käyttöön tiettyjä PCI-

laitteita, mutta tässä työssä pyritään pitämään konfiguraatio mahdollisimman yleisellä tasolla, joten näitä ominaisuuksia ei esitellä. Näin luodut virtuaalipalvelimet ovat helposti siirrettävissä toiselle isäntäkoneelle, joka ei välttämättä ole aivan identtinen.



Kuva 8 Kernel konfiguraatio

5.1.1 Asennus

Xenin asennnus Gentoo Linuxiin tapahtuu asentamalla kernel, josta löydyy xen-tuki, itse Xen-ohjelmisto sekä iSCSI-asiakasohjelmisto. Mitään muuta ei isäntäkoneen asennukseen tarvita.

5.1.2 Konfiguraatiot

Xend-taustaprosessin konfuguraatiotiedosto on oletusarvoisesti /etc/xend/xend-config.sxp. Tässä tiedostosssa määritellään esimerkiksi, missä verkoissa Xen kuuntelee virtuaalipalvelinten siirtopyyntöjä tai mihin verkkoihin isäntäkoneen virtuaalipalvelimet voivat liikennöidä. Tiedosto on Xen-isäntäkoneen kannalta yksi tärkeimmistä ja isäntäkoneen määrittelyt tehdään tässä tiedostossa. Kuvassa 8 nähdään Xen-isäntäkoneen ytimen konfiguraatiota.

Käytössä olevien asetusten selitykset:

- xend-relocation-port on portti, jossa xend kuuntelee migraatiopyyntöjä.
- xend-address on osoite, jossa xend HTTP-liittymä kuuntelee pyyntöjä.
- xend-relocation address on osoite, jossa xend kuuntelee migraatiopyyntöjä.
- xend-relocation-hosts-allow määrittelee osoitteet, joilta sallitaan migraatiopyynnöt.
- network-script määrittelee, mitä scriptiä kutsutaan kun xend luo uuden verkkoliittymän isäntäkoneelle.
- network-script network-bridge määrittelee verkkoyhteydeksi siltauksen.
- vif-script vif-bridge: skripti, jota kutsutaan, kun luodaan virtuaalipalvelimelle verkkoliityntä.
- dom0-min-mem määrittelee muistin määrän, jonka isäntäkone vähintään saa omaan käyttöönsä.
- dom0-cpus määrittelee kuinka montaa prosessoria (tai ydintä) isäntäkone voi käyttää omiin tarkoituksiinsa.

5.1.3 Virtuaalipalvelimen luonti

Helpoin tapa luoda virtuaalipalvelin on käyttää ns. levynkuvaa, johon luodaan tiedostojärjestelmä ja sitä käytetään kuin kovalevyä tai kovalevyn osiota. Sen sijaan tässä työssä käytetään LVM-pohjaista tiedostojärjestelmää iSCSI:n avulla verkosta, jolloin levyyn päästään käsiksi kaikilta ko. verkkoon kuuluvilta palvelimilta. Tämä mahdollistaa esimerkiksi virtuaalipalvelinten siirron isäntäkoneelta toiselle ilman katkosta palveluissa. Tiedostopohjainen virtuaalipalvelin on toisaalta helposti siirrettävissä palvelimelta toiselle, mutta suorituskyvyssä hävitään kuitenkin niin paljon, että se ei ole tuotantoympäristössä järkevää. Myös varmuuskopiot saadaan LVM:lla (Logical Volume Manager) toteutetusta järjestelmästä huomattavasti helpommin ja ennen kaikkea keskeytyksettä. Tässä kohtaa on myös hyvä huomata, että kun tiedostojärjestelmä ei ole tallennettuna paikallisesti isäntäkoneelle, niin myös tarvittavien kovalevyjen määrä vähenee ja näin myös virrankulutus ja lämmöntuotto

laskee. Puhumattakaan saavutetusta hyödystä – virtuaalipalvelinta voidaan ajaa miltä tahansa isäntäkoneelta.

Huomattavaa kuitenkin on, että Xen ei ’tiedä’, onko kyseinen virtuaalipalvelin jo käynnistetty toisella isäntäkoneella, joten käynnistyksen yhteydessä on oltava erityisen tarkka. Virtuaalipalvelinten ajo kiinteästi yhdellä isäntäkoneella poistaisi usemman käynnistyksen mahdollisuuden, mutta samalla menetettäisiin mahdollisuus siirtää virtuaalipalvelin katkotta jonkin toisen isäntäkoneen ajettavaksi. Ongelmia tulisi myös isäntäkoneen vikatilanteessa, jolloin virtuaalipalvelimia ei voida heti käynnistää toisella isäntäkoneella.

Graafinen tila ei ole saavutettavissa Xeinillä virtualisoiduissa virtuaalipalvelimissa kuten normaalissa Linux-järjestelmässä esimerkiksi xwindowsin avulla. Tämä on seurausta siitä, että virtuaalisessa järjestelmässä ei ole käytössä näytönohjainta, jolla kuva normaaliin tapaan näkyisi. Ratkaisuna on kuitenkin esimerkiksi muuttaa gdm:n (Gnome display manager) konfiguraatiota siten, että se käynnistääkin xwindowsin sijasta vnc-palvelimen ja näin päästään kiinni graafiseen tilaan kuten missä tahansa palvelimessa, jossa on vnc-palvelin asennettuna.

VNC:n käyttö onnistuu muuttamalla /etc/gdm/custom.conf -tiedostoa niin, että [servers] -osion alle muokataan gdm:n sijaan käynnistymään VNC-palvelin.

Tämä edellyttää, että xwindowsin ja Gnomen vaatimat ohjelmistot on asennettu onnistuneesti
.
Yksi hyvä ja varteenotettava vaihtoehto on myös NoMachine-nimisen yrityksen tuote FreeNX. Se on tarkoitettu verkkoihin, joissa siirtonopeus lähiverkkoja matalampi. FreeNX tarjoaakin huomattavasti mukavamman mahdollisuuden työskennellä työpöydän kanssa etäyhteyden avulla. (7.)

5.1.4 Palvelimen virtualisointi

Yksinkertainen Linux-Pohjainen sähköpostipalvein paravirtualisoidaan seuraavasti:

- Tehdään ensin tarvittavat levyosiot iSCSI-ryppäällä.
- Lisätään luodut levyalueet iSCSI-ryppään konfiguraatioon niin, että ne ovat saavutettavissa verkon avulla ja näin isäntäkoneiden käytettävissä.
- Uudelleenkäynnistetään iSCSI-taustaprosessi.

Nyt levyt ovat verkon avulla käytettävissä ns. iSCSI kohteina ja osoitteina on esimerkiksi iqn.2007-01.fi.tentacle:storage.iscsi01.tentacle.mailroot, jolla ne myöhemmin löydetään.

Haluttu levyosio saadaan käyttöön Xen-isäntäkoneella iscsiadm-komennon avulla. Levyalueiden käyttöönotto tulee suorittaa kaikilla isäntäkoneilla jokaista levyaluetta kohti, joita virtuaalipalvelimet tai isäntäkone itse tarvitsevat.

Verkon ja iSCSI:n avulla käyttöön otetulle levyalueelle luodaan tiedostojärjestelmä ja se otetaan isäntäkoneella käyttöön mount-komennolla. Tämän jälkeen virtualisoitavan Linux-palvelimen sisältö voidaan siirtää verkon avulla rsync-komennolla. Kopiointi voidaan suorittaa lähdekoneen ollessa täysin toiminnassa ja ilman, että taustalla menevää kopiointia edes käyttäjät huomaisivat. Kannattaa katsoa, että rsync-komennolle annetaan oikeat parametrit, jotta tiedostojen omistajuus ja oikeudet säilyvät haluttuina. Tämä on ensisijaisen tärkeää niin Linuxin itsensä toiminnalle kuin palvelimen tietoturvallekin.

Kopioinnin jälkeen järjestelmään tehdään muutamia muutoksia chroot-komennon avulla. Tärkeimmät muutokset ovat Linuxin käynnistyksen kannalta oleellisten tiedostojen luonti. Aivan kaikki tiedostot eivät kuitenkaan rsync-komennolla siirry. Kopioinnin jälkeen uotavat tiedostot ovat /dev/null, /dev/zero, /dev/console ja /dev/tty1.

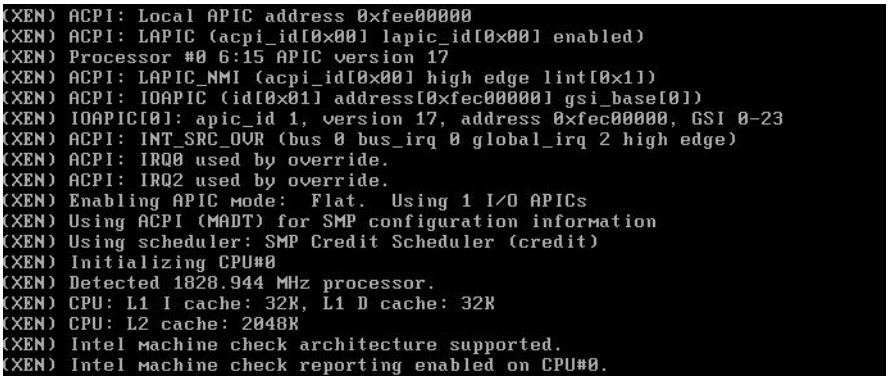
Tämän jälkeen virtuaalipalvelimen tiedostojärjestelmä tiedostoineen on valmiina käyttöönottoon.

Seuraavaksi luodaan virtuaalipalvelimelle konfiguraatiotiedosto, jossa kerrotaan sen ominaisuudet kuten nimi, muistin määrä, levyalueet ja se, mitä vikatilanteessa tehdään. Konfiguraatiot sijoitetaan verkon avulla jaetulle levyalueelle, joka on näkyvissä kaikille isäntäkoneille. Kaikkien isäntäkoneiden tulee nähdä kyseinen levyalue siksi, että virhetilanteessa virtuaalipalvelimia voidaan käynnistää miltä tahansa isäntäkoneelta.

Virtuaalipalvelimen konfiguraatiosta löytyy esimerkiksi seuraavat optiot:

- name määrittelee nimen, jolla virtuaalipalvelin näkyy käytettäessa xenin tarjoamia hallintatyökaluja kuten ’xm’.
- memory allokoi virtuaalipalvelimelle muistia.
- kernel: osoittaa Linux-ytimen, jolla virtuaalipalvelin käynnistetään.
- disk osoittaa tarvittavat levyalueet ja niiden nimet kirjoitusoikeuksineen
- root osoittaa root-tiedostojärjestelmän.
- vif ottaa verkon käyttöön virtuaalipalvelimelle. Tässä voitaisiin määritellä myös verkkoliityntä tai MAC-osoite, jota virtuaalipalvelimen tulee käyttää.
- vcpus allokoi virtuaalipalvelimelle prosessorit tai ytimet.
- on_crash määrittelee, että vikaantuessa virtuaalipalvelin uudelleenkäynnistetään.
- on_reboot määrittelee, että uudelleenkäynnistettäessä palvelin käynnistetään uudelleen normaalisti.
- on_poweroff määrittelee, että sammutettaessa palvelinta ei enää uudelleenkäynnistetä vaan se pitää täytyy manuaalisesti.

Tämän jälkeen virtuaalipalvelin on valmis käynnistettäväksi. Palvelinta käynnistettäessä ensimmäistä kertaa tai virhettä selvitettäessä kannattaa käyttää xm:n optiota ’console’. Optio määrittelee, että palvelimeen yhdistetään heti ns. virtuaalinen konsoliyhteys, jolla nähdään käynnistyksessä mahdollisesti esiintyvät ongelmat kuten normaalisti palvelimen näytöltä. Virtuaalipalvelimen käynnistys voidaan nähdä kuvasta 9.



Kuva 9 Xen virtuaalipalvelimen käynnistys

Xen tarjoaa myös mahdollisuuden säädellä virtuaalipalvelimen käytössä olevaa muistia sen ollessa käynnissä. Maksimimuistin määrä tulee kuitenkin asettaa konfiguraatio-tiedostoon parametrilla maxmem, jotta tämä piirre saadaan käyttöön. Muistia voidaan sitten lisätä tai vähentää komennolla xm mem-set. Tämä on erittäin hyödyllinen esimerkiksi seuraavissa tapauksissa:

- Virtuaalipalvelin kärsii suorituskykyongelmista ja liiallisesta swap-muistin käytöstä
- Halutaan käynnistää lisää prosesseja, jotka vaativat muistin lisäämistä
- Muistia jää yli ja se voidaan allokoida muille virtuaalipalvelimille. (5.)

5.2 Hallinta ja konsolityöskentely

Xen-virtuaalipalvelinten hallintaan on tarjolla melko laaja valikoima vapaita työkaluja. Esittelen tässä lyhyesti näistä muutamia mielenkiintoisimmista.

5.2.1 Konsoli

Virtuaalipalvelinten hallintaan parhaat perustyökalut löytyvät konsolista. Käyttömukavuus ei ehkä vastaa kaikkien mieltymyksiä, mutta näiden tunteminen on välttämätöntä tilanteessa, jossa graafiseen ympäristöön ei syystä tai toisesta ole pääsyä. Hyvin harvassa palvelimessa myöskään on tapana asentaa graafista ympäristöä. Keskeisin työkalu, jolla Xen-virtuaalipalvelimia hallitaan ja tarkastellaan on komento nimeltään *xm* (xen management). Alla muutamia yleisiä komentoja, joita voidaan käyttää isäntäkoneelta virtuaalipalvelinten hallintaan:

xm list listaa isäntäkoneella ajossa olevat virtuaalipalvelimet ja antaa niistä yleistä tietoa.

Taulukossa 2 on esimerkki *xm list* -komennon käytöstä:

Taulukko 2 xm list -komennon käyttö

```
virtual-external-01 ~ # xm list
Name                        ID   Mem VCPUs      State   Time(s)
Domain-0                     0  1046     2          r-----  45448.0
ldap                        12   256     1          -b----   6640.9
ns1                         19   256     1          -b----   5470.9
shell                       10   512     1          -b----  17490.6
www                         18  1024     2          -b----  29536.8
```

Taulukossa 2 nähdään xm-komennon tuottama listaus. Tulosteen selitykset ovat seuraavat:

- Name: virtuaalipalvelimen nimi, joka määriteltiin konfiguraatiotiedostossa.
- ID: virtuaalipalvelimen id. Sammutettaessa virtuaalikone se menettää id:n ja se luodaan uudelleen taas käynnistyksen yhteydessä.
- Mem: virtuaalipalvelimelle allokoidun muistin määrä
- VCPUs: virtuaalipalvelimelle allokoitujen prosessorien määrä.
- State: virtuaalipalvelimen tila, joka voi olla b (blocked), c (crashed), p (paused), r (running) tai s (shutdown). Blocked -tila voisi olla hyvin nimetty myös 'idle', koska kun virtuaalipalvelin ei tarvitse prosessoriaikaa sen tilaksi merkitään b. Tämän sarakkeen tulkinta aiheuttaa usein väärinkäsityksiä.
- Time(s): Prosessoriaika, jonka virtuaalipalvelin on käyttänyt ajossa ollessaan.

Muita xm-komennon käytetyimpiä optioita ovat:

xm create /konfiguraation/polku - käynnistää virtuaalipalvelimen.
xm reboot - uudelleenkäynnistää virtuaalipalvelimen.
xm shutdown - sammuttaa virtuaalipalvelimen.
xm destroy - välittömästi sammuttaa virtuaalipalvelimen.
xm console - antaa konsoliyhteyden virtuaalipalvelimelle.
xm mem-set - lisää tai poistaa muistia virtuaalipalvelta.
xm migrate - siirtää virtuaalipalvelimen toiselle isännälle.
xm pause - pysäyttää virtuaalipalvelimen ajon.
xm resume - palauttaa pysäytetyn virtuaalipalvelimen ajoon.

Komento *xentop* vastaa Linuxista ja Unixista tuttua *top*-komentoa, mutta näyttää sen sijaan tietoa virtuaalipalvelimista ja isäntäkoneesta, niiden verkkoyhteyksistä, prosessorinkäytöstä ja niin edelleen. Tässä onkin huomattavaa, että vaikka isäntäkoneen *top*-komento näyttää esimerkiksi tilanteen, jossa CPU:ta on käytössä 0 % niin totuus tulee ilmi kuitenkin *xentop*-komennolla, joka voi hyvinkin olla esimerkiksi 99 %. *Top*-komennon sijaan Xen-isäntäkoneilla tuleekin aina käyttää *xentop*-komentoa jos halutaan seurata virtuaalipalvelinten tilaa.

5.2.2 Valvonta

Virtuaalipalvelinten ja isäntäkoneiden valvontaan käytetään ohjelmistoa nimeltään Nagios. Nagios valittiin valvontatyökaluksi sen erinomaisen skriptattavuutensa ja yksinkertaisuutensa vuoksi. Tämän lisäksi Nagiokseen on saatavilla vielä liitännäinen, joka valvoo tiettyjen virtuaalipalvelinten tilaa ja hälyttää vikatilanteista. Nagios siis seuraa valvottaviksi annettujen palvelinkoneiden tai palveluiden tilaa ja hälyttää tarvittaessa ennaltamäärätyin keinoin, kuten sähköposti tai tekstiviesti. Koska aikaisemmin todettiin, että virtuaalipalvelimia ei pystytä tällä hetkellä luotettavasti automaattisesti siirtämään toisen isäntäkoneen ajetttaviksi, niin hyvinkin alkeelliset valvontametodit tulivat kyseeseen. Valvonassa käytetään liitännäistä joka tarkastaa, että tietyt virtuaalipalvelimet ovat ajossa ja että isäntäköne on toiminnassa. Jos näin ei ole, lähettää valvontapalvelin asiasta sähköpostin ja tekstiviestin kahdelle yrityksen teknisestä kapasiteetista vastaavalle henkilölle. Kuvissa 10 ja 11 on Nagioksen valvontaruudut, joita voidaan seurata selaimella.

tentacle-customer-01-xen	PING	OK	02-15-2008 02:28:06	6d 10h 49m 28s	1/5	PING OK - Packet loss = 0%, RTA = 5.33 ms
	Xen Virtual Machine Monitor	OK	02-15-2008 02:24:38	28d 15h 22m 1s	1/4	OK: Xen Hypervisor "virtual-customer-01" is running Xen VMs: customer001
tentacle-external-01-xen	PING	OK	02-15-2008 02:28:06	0d 3h 58m 48s	1/5	PING OK - Packet loss = 0%, RTA = 2.14 ms
	Xen Virtual Machine Monitor	OK	02-15-2008 02:28:38	2d 22h 57m 59s	1/4	OK: Xen Hypervisor "virtual-external-01" is running Xen VMs: mail
tentacle-external-02-xen	PING	OK	02-15-2008 02:28:06	6d 10h 49m 28s	1/5	PING OK - Packet loss = 0%, RTA = 11.80 ms
	Xen Virtual Machine Monitor	OK	02-15-2008 02:27:54	26d 9h 9m 29s	1/4	OK: Xen Hypervisor "virtual-external-02" is running Xen VMs: ldap

Kuva 10 Nagios-valvonta Xen-virtuaalipalvelimilta

mail	/	OK	02-15-2008 02:28:03	2d 21h 11m 0s	1/4	DISK OK - free space: / 7041 MB (61% inode=77%):
	/var/vpopmail	OK	02-15-2008 02:29:06	2d 21h 10m 12s	1/4	DISK OK - free space: /var/vpopmail 11999 MB (78% inode=-):
	Load	OK	02-15-2008 02:27:06	1d 6h 59m 53s	1/4	OK - load average: 0.43, 0.28, 0.28
	PING	OK	02-15-2008 02:31:06	7d 2h 39m 7s	1/5	PING OK - Packet loss = 0%, RTA = 13.95 ms
	SMTP	OK	02-15-2008 02:28:10	0d 6h 43m 46s	1/4	SMTP OK - 1.092 sec. response time
	Total Processes	OK	02-15-2008 02:27:07	2d 21h 10m 8s	1/4	PROCS OK: 131 processes

Kuva 11 Nagios prosessienvalvonta

5.2.3 Migraatio

Manuaalinen migraatioprosessi on tiedostopohjaisella virtuaalipalvelimella helppo, mutta palveluiden kannalta erittäin huono vaihtoehto. Manuaalisella migraatiolla tarkoitetaan yksinkertaisesti virtuaalipalvelimen levynkuvan ja konfiguraation kopiointia isäntäkoneelta toiselle. Virtuaalipalvelin joudutaan pysäyttämään kopioinnin ajaksi, mutta tämän jälkeen virtuaalikone voitaisiin käynnistää normaalisti. Manuaalinen migraatio on helppo prosessi, mutta yleisesti ottaen se soveltuu vain Xen-virtualisointiin tutustumiseen tai kokeilupalvelimille. Syynä on yksinkertaisesti operaation hitaus, koska isojen tiedostojen ollessa kyseessä nopeakaan verkko ei ole riittävän nopea suoriutuakseen kopioinnista niin nopeasti, että sitä voitaisiin pitää yritysmaailmassa vaihtoehtona. Vielä kopioinnin jälkeenkin tulisi vielä odottaa virtuaalipalvelimen käynnistymistä, joka virtuaalipalvelimesta riippuen voi viedä kauankin.

Xen-virtuaalipalvelimia pystytään myös siirtämään verkon avulla isäntäkoneelta toiselle käytännössä katsoen ilman katkoa palvelussa. Tätä kutsutaankin nimellä ’live migraatio’. Tässä prosessissa virtuaalipalvelimen muistin sisältö kopioidaan kohteena olevalle isäntäkoneelle ilman, että täysin pysäytettäisiin siirrettävänä oleva virtuaalipalvelin ja käynnistettäisiin uudella isäntäkoneella. Siirrosta aiheutuu hyvin pieni katko, luokkaa 60–300 ms. Huomattavaa kuitenkin on, että tämä onnistuu vain, jos käytössä on jaettu levyalue, kuten esimerkiksi NFS tai kuten tässä tapauksessa iSCSI, jonka isäntäkoneet voivat halutessaan ottaa käyttöön. Tiedostopohjaisilla virtuaalipalvelimilla tämä ei ole mahdollista. (1.)

Migraation eteneminen on nähtävissä kuvassa 13.

Kuva 13 (4.)

*1. Ennen migraatiota***:** Virtuaalipalvelin (kuvassa domU) on normaalisti käynnissä isäntäkoneella. Virtuaalipalvelin on aktiivisessa tilassa.

2. Resurssien varaaminen: Migraatiopyyntö lähetetään uudelle isäntäkoneelle, joka tarkistaa ja varaa tarvittavat resurssit. Jos uuden isäntäkoneen hypervisor ei pysty varaamaan tarpeeksi resursseja migraation toteuttamiseen prosessi keskeytetään. Migraation keskeytyessä virtuaalipalvelin jatkaa toimintaansa normaaliin tapaan vanhalla isäntäkoneellaan.

3. Aloitetaan muistin kopiointi: Virtuaalipalvelimen käytössä olleen muistin kopiointi uudelle isäntäkoneelle aloitetaan. Aluksi lähdetään kopioimaan koko muistialue ja lopuksi tarkistetaan muuttuneet osat, jotka kopioidaan viimeisenä.
4. Pysätys ja kopiointi: Aktiivinen ja normaalissa toiminnassa oleva virtuaalipalvelin pakotetaan pysäytettyyn tilaan ja sen verkkoliikenne ohjataan uudelle isäntäkoneelle. Tämän vaiheen lopuksi uudella ja vanhalla isäntäkoneella on suspended-tilassa oleva kopio virtuaalipalvelimesta. Vanhalla isäntäkoneella oleva kopio on vielä ns. pääkopio. Migraation päättyminen virheeseen palauttaa tämän kopion toimintaan.

5. Vapautetaan virtuaalipalvelin: Uusi isäntäkone lähettää vanhalle kuittauksen, että sillä on täydellisesti kopioitu virtuaalipalvelin ajossaan. Vanha isäntäkone poistaa virtuaalipalvelimen ja uusi isäntäkone ottaa sen pysyvästi ajettavakseen.

6. Virtuaalipalvelin aktivoidaan: Uusi isäntäkone aktivoi virtuaalipalvelimen. Kaikki virtuaalipalvelimen sisäiset ajurit yhdistetään uudelleen ja se palautuu normaaliin ajoon. (4.)

5.2.4 Varmistus

Varmistuksissa, samoin kuin migraatiossa helpoimmalla pääsee, jos virtuaalipalvelimet ovat tiedostopohjaisia ja kestävät olla hetken aikaa tavoittamattomissa. Tällöin virtuaalipalvelin voidaan vain pysäyttää ja kopioda palvelimen sisältävä tiedosto talteen. Tämän jälkeen virtuaalipalvelin vain käynnistettäisiin normaalisti uudelleen. Myös siirto toiselle isäntäkoneelle tällä tavoin on erittäin kätevää, koska tarvitaan vain konfiguraatiotiedosto ja tiedosto, joka sisältää virtuaalipalvelimen tiedostojärjestelmän. Tämä on harvoin kuitenkaan tuotantoympäristössä mahdollista, joten varmistukset tehdään LVM:n tilannevedos-ominaisuudella. Tilannevedos mahdollistaa tiedostojärjestelmän kopioinnin tausta-ajona keskeyttämättä kopioitavaa virtuaalipalvelinta. Käytännössä prosessi toimii niin, että ensin iSCSI-ryppäällä luodaan tilannevedos halutusta virtuaalipalvelimen levyosiosta ja luotu tilannevedos voidaan ottaa käyttöön kuten normaali levyalue ja kopioda sen sisältö rsync -komennolla varmistuspalvelimelle, josta palvelimen tiedostot ovat nopeasti palautettavissa verkon avulla virhetilanteen sattuessa. Varmistuspalvelimella säilytetään jokaisen virtuaalikoneen data seitsemältä päivältä taaksepäin. Kerran viikossa jokainen virtuaalipalvelin myös tallennetaan nauharobotille. Nauhalla varmistuksia säilytetään neljä kappaletta eli kuukauden ajan.

5.2.5 Palautus

Virtuaalipalvelimen palautusprosessi on tiedostopohjaisella palvelimella hyvin yksinkertaista. On vain palautettava palvelimen sisältävä tiedosto ja käynnistettävä palvelin uudelleen. Tässä työssä virtuaalipalvelimet on varmistettu LVM:n snapshot-ominaisuudella. Palautusprosessi on siis astetta vaikeampi, mutta kuitenkin erittäin suoraviivainen. Ensin luodaan iSCSI-ryppäällä virtuaalipalvelinta varten tarvittava levyosio. Katsotaan, että kapasiteetti varmasti riittää palautukselle. Seuraavaksi kyseinen levyalue jaetaan iSCSI:n kautta verkkoon, josta se saadaan Xen-isäntäkoneiden käyttöön. Tämän jälkeen levy alustetaan ja kopioidaan varmistuksista kyseisen virtuaalipalvelimen sisältö varmistuksista esimerkiksi rsync-komennolla.

Kopioinnin jälkeen virtuaalipalvelin on valmis käynnistettäväksi normaaliin tapaan. Toipumisprosessi on siis vain minuutteja kestävä operaatio, jota voidaa pitää huomattavasti nopeampana kuin fyysisen palvelimen tiedostojen palautusta ja laitteiston konfigurointia.

6 Yhteenveto

Tässä työssä tutkittiin Linux-jakeluiden paravirtualisointia Xen-virtualisointialustalla. Työn kirjallisessa osuudessa tarkasteltiin, miten Xen-virtualisointilusta teoriassa toimii ja vertailtiin tekniikoita keskenään. Todettiin, että Xen sopii yrityksen tarpeisiin parhaiten, koska merkittävässä roolissa olivat eri ratkaisujen kustannukset. Lisäksi paravirtualisoidut Linux-jakelut ovat erittäin suorituskykyisiä.

Asiakkaiden ja yrityksen oman henkilökunnankin kannalta vasteaika ja saatavuus ovat palveluiden oleellisia laatumittareita. Saatavuuteen ja vasteaikoihin pyrittiin vaikuttamaan niitä parantavasti uudella vikasietoisella ympäristöllä.

Käytännön osuudessa luotiin vikasietoinen palvelinympäristö, jossa pyrittiin eri kompnentit ja fyysiset palvelimet kahdentamaan. Minkään yksittäisen komponentin tai fyysisen palvelimen vikatilanne ei aiheuta kuin lyhyen katkoksen palveluiden toimintaan.

Käytännössä ympäristöä seuranneena voidaan todeta, että Xen-paravirtualisointi oli oikea ratkaisu ja ympäristö on toiminut moitteetta käyttöönottonsa jälkeen. Myös kuluja saatiin pienennettyä palvelinmäärän vähenemisen myötä.

Lähteet

1 Xen wikipedia. (WWW-dokumentti.) <http://en.wikipedia.org/wiki/Xen>. Lainattu 24.10.2007

2 Habib, Irfan. Xen. Linux Journal 30.5.2006

3 Williams, David. Virtualization with Xen. Massachusetts: Syngress Publishing Inc., 2007.

4 Chaganti, Prabhakar. Xen Virtualization. Birmingham: Packt Publishing, 2007.

5 Von Hagen, William. Professional Xen Virtualization. Indianapolis: Wiley Publishing Inc., 2008.

6 Wolf, Chris., Halter, Eric. Virtualization from the Desktop to the Enterprise. California: Apress, 2005.

7 Matthews, Jeanna., Dow, Eli., Deshane, Todd., Hu, Wenjin., Bongio, Jeremy., Wilbur, Patric., Johnson, Brendan. Running Xen a Hands-on Guide to the Art of Virtualization. Massachusetts, 2008.

8 Wikipedia. (WWW-dokumentti.) <http://en.wikipedia.org/wiki/Thin_client>. Lainattu 16.9.2010

9 Citrix-verkkosivusto. (WWW-dokumentti.) <http://www.citrix.com/virtualization/desktop-virtualization.html>. Lainattu 17.10.2010